Attitude Estimation Based on Solution of System of Polynomials via Homotopy Continuation

Yang Cheng*

Mississippi State University, Mississippi State, MS 39762

John L. Crassidis†

University at Buffalo, State University of New York, Amherst, NY 14260-4400

The optimal attitude estimate is often defined as the solution to a minimization problem. When the objective function of the minimization problem is quadratic in the attitude matrix or equivalently quartic in the attitude quaternion, for example, in GPS attitude determination, gradient-based iterative algorithms are usually used, which can only find the local minimizer. A novel numerical method is presented that is guaranteed to find the global minimizer of the minimization problem. It first finds all the stationary points of the minimization problem by solving the polynomial equations satisfied by the stationary points using the homotopy continuation based solver and then chooses the global minimizer from them. Two numerical examples are used to show the viability of the method for attitude estimation.

I. Introduction

Attitude estimation is the process of determining the orientation of the body frame of a vehicle with respect to some reference frame. The attitude is not directly measured but needs to be inferred from directional or angular measurements. Given a set of vector and/or direction cosine observations, the optimal attitude estimate is usually determined as the solution to a minimization problem, formulated based on the least-squares[1] or maximum likelihood estimation principle.[2] The resultant objective function of the minimization problem is quadratic in the residuals. The minimization problem is subject to one or more equality constraints if redundant attitude representations such as the attitude quaternion and the attitude matrix are used. For example, the attitude quaternion needs to be of unity norm and the attitude matrix needs to be a proper orthogonal matrix.[3] Three-parameter attitude parameters such as the Euler angles, Rodrigues parameters, and modified Rodrigues parameters are not subject to constraints, but are singular or discontinuous at some attitudes.[3] They are therefore not as widely used in three-axis attitude estimation as the attitude quaternion or the attitude matrix.

Example minimization problems for attitude estimation include the Wahba problem[1,2,4–6] and the GPS attitude determination problem (with resolved integer ambiguity).[7–9] The former determines the attitude from vector observations; the latter determine the attitude from direction cosine measurements. The Wahba problem uses special weighting matrices in the objective function. As a result, its objective function only linearly depends on the attitude matrix and it has closed-form solutions based on the singular value decomposition[5] or eigenvalue/eigenvector decomposition.[4] The GPS attitude determination problem, whose objective function is quadratic in the attitude matrix and quartic in the attitude quaternion except for very special cases, does not have closed-form exact solutions. Closed-form solutions for a transformed problem that is equivalent to the Wahba problem are possible.[7] Gradient-based iterative methods are usually used to solve such problems.[8,9] However, these methods are not guaranteed to find the global minimum and their performances depend on the initial attitude guess.

This paper presents a novel method for the minimization problems whose irreducible objective functions are quadratic in the attitude matrix or equivalently quartic in the attitude quaternion. The basic idea of the

*Assistant Professor, Department of Aerospace Engineering. Email: cheng@ae.msstate.edu. Associate Fellow AIAA.
†Professor, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Associate Fellow AIAA.

method is to 1) find all the stationary points of the minimization problem and 2) choose the global minimizer from them as the one with the smallest objective function. The stationary points are the points at which the gradient of the objective function vanishes. For minimization problems with equality constraints only, the global minimizer must be one of the stationary points.[10] The solution thus obtained is guaranteed to be globally optimal. By computing the Hessian matrix (the second derivative of the objective function), the method can find all the local minimizers as well, which are the stationary points with positive definite Hessian matrix.[10] While computing the Hessian matrices and finding all the local minimizers are not essential to picking up the global minimizer from the stationary points, which only requires evaluation of the objective function, knowledge about the number of local minimizers and their locations helps to understand the properties of the minimization problem.

The equations the stationary points satisfy are or can be transformed into a small-scale low-degree system of polynomials. The number of equations is $2n-3$, where n is the number of parameters of an attitude representation. For example, the polynomial system for the attitude quaternion has five equations, of which the highest degree is three; the polynomial system for the attitude matrix has 15 equations, of which the highest degree is two. These small-scale low-degree polynomial equations can be reliably solved using homotopy continuation methods.[11,12] It should be noted that the homotopy continuation method finds all isolated complex roots of a polynomial system instead of real roots only. Since only real roots have physical meaning, post-processing is need to eliminate the complex roots with non-zero imaginary part. There is no general method to find all the real roots directly, however. Several tools for solving systems of polynomials by homotopy continuation exist, such as HOM4PS,[13] Bertini,[14] PHCpack[15] and HomLab.[12] According to our tests, HOM4PS is the fastest and thus the tool of choice. It can solve the polynomial systems in the paper in less than a second (see the simulation results).

The contribution of the paper is the application of the homotopy continuation method to attitude estimation. The paper is primarily concerned with how to derive the polynomial systems in the attitude estimation problem, but not how to solve the polynomial systems.

The organization of the rest of the paper proceeds as follows. First, the observation models, the Wahba problem, and the GPS attitude determination problem are reviewed. Then, a general minimization problem for attitude estimation is formulated. Next, the quaternion and attitude matrix based solutions to the minimization problem are presented. Finally, simulation results of the Wahba problem and the GPS attitude determination problem are given, followed by the conclusions.

II. Minimization Problems in Attitude Estimation

A. Observation Models

Many types of attitude measurement data can be converted to unit vector and direction cosine observations. The unit vector and direction cosine observation models are respectively given by[4,7]

$$\mathbf{b}_i = A\mathbf{r}_i + \boldsymbol{\upsilon}_i \tag{1}$$

$$d_{ij} = (\mathbf{b}_i^d)^T A \mathbf{r}_j^d + \upsilon_{ij} \tag{2}$$

where A is the attitude matrix, the superscript T denotes the transpose of a matrix or vector, $\mathbf{b}_i$ and $\mathbf{b}_i^d$ are the representations of unit vectors in the body frame, $\mathbf{r}_i$ and $\mathbf{r}_i^d$ the representations of unit vectors in the reference frame, $\boldsymbol{\upsilon}_i$ and υ_{ij} are the noise terms. The vectors are assumed to be of unity length because the magnitude of the vectors in the models does not contain attitude information. The superscript d in $\mathbf{b}_i^d$ and $\mathbf{r}_i^d$ means that the vectors are used in the directional cosine model. The vectors $\mathbf{b}_i$ and $\mathbf{b}_i^d$ (or $\mathbf{r}_k$ and $\mathbf{r}_i^d$) are not necessarily identical. The attitude matrix is also known as the direction cosine matrix or the rotation matrix. It is a proper orthogonal matrix, i.e., $A^TA = AA^T = I_{3\times 3}$, $\det A = +1$, where det denotes the matrix determinant.[3] The noise vectors in the vector observation model are assumed to be zero-mean and covariance R_i (to first order):

$$E[\boldsymbol{\upsilon}_i] = \mathbf{0}_{3\times 1} \tag{3a}$$

$$E[\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_i^T] = R_i \tag{3b}$$

Because $\mathbf{b}_i$ and $\mathbf{r}_i$ are both assumed to be unit vectors, the covariance matrix R_i is singular (to first order). The null vector of R_i is $(A\mathbf{r}_i)$:

$$R_i(A\mathbf{r}_i) = \mathbf{0}_{3\times 1} \tag{4}$$

The noise in the direction cosine observation model is assumed to be zero-mean and variance σ_{ij}^2:

$$E[v_{ij}] = 0 \tag{5a}$$
$$E[v_{ij}^2] = \sigma_{ij}^2 \tag{5b}$$

Maximum likelihood estimation is widely used in attitude estimation.[16] To apply maximum likelihood estimation, the probability distributions of the noise terms need to be known. They are usually assumed to be Gaussian. Under the Gaussian assumption, it can be shown that the log-likelihood functions for the two observation models are quadratic in $(\mathbf{b}_i - A\mathbf{r}_i)$ and $(d_{ij} - (\mathbf{b}_i^d)^T A\mathbf{r}_j^d)$, respectively. The optimal attitude estimate in the maximum likelihood sense is therefore the solution to the optimization problem with objective function quadratic in the residuals $(\mathbf{b}_i - A\mathbf{r}_i)$ or $(d_{ij} - (\mathbf{b}_i^d)^T A\mathbf{r}_j^d)$.

B. Wahba Problem

The Wahba problem is to determine the optimal attitude from vector observations. The objective function is of the form

$$J(A) = \frac{1}{2}\sum_{i=1}^{n_v} (\mathbf{b}_i - A\mathbf{r}_i)^T W_i (\mathbf{b}_i - A\mathbf{r}_i) \tag{6}$$

where W_i are the weighting matrices and n_v is the number of vector observations. The expanded form of the objective function is given by

$$J(A) = \frac{1}{2}\sum_{i=1}^{n_v} \left(\mathbf{r}_i^T A^T W_i A\mathbf{r}_i - 2\mathbf{b}_i^T W_i A\mathbf{r}_i + \mathbf{b}_i^T W_i \mathbf{b}_i\right) \tag{7}$$

with the quadratic term in A $\mathbf{r}_i^T A^T W_i A\mathbf{r}_i$. The Wahba problem makes an assumption about the weighting matrices W_i that leads to the closed-form solution of the problem. It is assumed that W_i are the 3×3 identity matrix up to a scalar, given by

$$W_i = \frac{1}{\sigma_i^2} I_{3\times 3} \tag{8}$$

where the variance σ_i^2 is the single parameter used to characterize the statistical properties of $\boldsymbol{v}_i$ in the unit vector observation model. With this special form of W_i,

$$\mathbf{r}_i^T A^T W_i A\mathbf{r}_i = \frac{\mathbf{r}_i^T A^T A\mathbf{r}_i}{\sigma_i^2} = \frac{\mathbf{r}_i^T \mathbf{r}_i}{\sigma_i^2} = \frac{1}{\sigma_i^2} \tag{9}$$

That is, $\mathbf{r}_i^T A^T W_i A\mathbf{r}_i$ no longer depend on A. The objective function is now linear in A and can be written as[4, 6]

$$\begin{aligned} J(A) &= \sum_{i=1}^{n_v} \frac{1}{\sigma_i^2}\left(1 - \mathbf{b}_i^T A\mathbf{r}_i\right) \\ &= \sum_{i=1}^{n_v} \frac{1}{\sigma_i^2} - \mathrm{trace}[AB^T] \end{aligned} \tag{10}$$

where the 3×3 B matrix is known as the attitude profile matrix, given by[4, 6]

$$B = \sum_{i=1}^{n_v} \frac{1}{\sigma_i^2}\mathbf{b}_i\mathbf{r}_i^T \tag{11}$$

A closed-form solution to the Wahba problem is based on the singular value decomposition of B, given by

$$B = USV^T \tag{12}$$

where S is a diagonal matrix, and U and V are orthogonal matrices satisfying

$$UU^T = U^T U = I_{3\times 3} \tag{13a}$$
$$VV^T = V^T V = I_{3\times 3} \tag{13b}$$

The optimal attitude $\hat{A}$ is then given by[5]

$$\hat{A} = U \operatorname{diag}([1, 1, d]) V^T \tag{14}$$

$$d = (\det U)(\det V) \tag{15}$$

This singular value decomposition solution is considered closed-form because robust and accurate general-purpose numerical routines are available for the singular value decomposition.

Another closed-form solution to the Wahba problem is based on the quaternion parameterization of the attitude matrix and eigenvalue/eigenvector decomposition.[4] The attitude quaternion $\mathbf{q} = [q_1, q_2, q_3, q_4]^T$ is related to the attitude matrix by[3]

$$A = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2q_1q_2 + 2q_3q_4 & 2q_1q_3 - 2q_2q_4 \\ 2q_1q_2 - 2q_3q_4 & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2q_1q_4 + 2q_2q_3 \\ 2q_1q_3 + 2q_2q_4 & 2q_2q_3 - 2q_1q_4 & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \tag{16}$$

The attitude quaternion always satisfies $\mathbf{q}^T\mathbf{q} = q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1$ and one attitude matrix always corresponds to two attitude quaternions, $\pm\mathbf{q}$. Because of the unity-length constraint of the attitude quaternion, the expression for the attitude matrix in terms of the attitude quaternion is not unique. For example,

$$A = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2q_1q_2 + 2q_3q_4 & 2q_1q_3 - 2q_2q_4 \\ 2q_1q_2 - 2q_3q_4 & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2q_1q_4 + 2q_2q_3 \\ 2q_1q_3 + 2q_2q_4 & 2q_2q_3 - 2q_1q_4 & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} + (q_1^2 + q_2^2 + q_3^2 + q_4^2 - 1)^l L(\mathbf{q}) \tag{17}$$

with $l \geq 1$ an integer and L an arbitrary function of $\mathbf{q}$ is identical to Eq. (16).

Because the objective function is linear in A and A is quadratic in $\mathbf{q}$, the objective function is equivalent to a quadratic form in $\mathbf{q}$, given by[4,6]

$$J(\mathbf{q}) = J(A(\mathbf{q})) = \sum_{i=1}^{n_v} \frac{1}{\sigma_i^2} - \mathbf{q}^T K \mathbf{q} \tag{18}$$

with

$$K = \begin{bmatrix} B + B^T - \operatorname{trace}[B] I_{3\times 3} & \mathbf{z} \\ \mathbf{z}^T & \operatorname{trace}[B] \end{bmatrix} \tag{19}$$

where

$$\mathbf{z} = \begin{bmatrix} B_{23} - B_{32} \\ B_{31} - B_{13} \\ B_{12} - B_{21} \end{bmatrix} \tag{20}$$

The optimal attitude quaternion $\hat{\mathbf{q}}$ is the eigenvector corresponding to the maximum eigenvalue of the 4×4 K matrix.[4,6]

C. GPS Attitude Determination

GPS attitude determination is based on the direction cosine observation model. It is assumed that the integer ambiguities in the carrier wave phase measurements have been resolved. Compared with the Wahba problem, GPS attitude determination does not have closed-form solutions except for very special cases that reduce to the Wahba problem. The objective function for GPS attitude determination is given by[7]

$$J(A) = \frac{1}{2} \sum_{i=1}^{m_d} \sum_{j=1}^{n_d} w_{ij} \left[d_{ij} - (\mathbf{b}_i^d)^T A \mathbf{r}_j^d \right]^2 \tag{21}$$

where $w_{ij} = 1/\sigma_{ij}^2$ and m_d is the number of baselines $\mathbf{b}_i^d$ and n_d the number of GPS sightlines $\mathbf{r}_i^d$. The quadratic term in A is $\left((\mathbf{b}_i^d)^T A \mathbf{r}_j^d\right)^2$. In general, the problem is solved using gradient-based method, which depends on an initial guess.[8,9]

D. Weighted Orthogonal Procrustes Problem

Both the Wahba problem and the GPS attitude determination problem are related to the weighted orthogonal Procrustes problem. The objective of the problem is to find an orthogonal matrix that minimizes the following objective function:[17]

$$J(A) = \frac{1}{2} \|\mathcal{B} A \mathcal{C} - \mathcal{D}\|_F^2 \tag{22}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix and $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are matrices of appropriate dimensions. Note that $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are not necessarily square matrices. While the orthogonal Procrustes problem is not limited to three-dimensional orthogonal matrices, the orthogonal matrix in this paper is always three-dimensional with determinant $+1$, that is, an attitude matrix.

The objective function can be rewritten as

$$\begin{aligned} J(A) &= \frac{1}{2} \operatorname{trace}\left[(\mathcal{B} A \mathcal{C} - \mathcal{D})(\mathcal{B} A \mathcal{C} - \mathcal{D})^T\right] \\ &= \frac{1}{2} \operatorname{trace}\left[A \mathcal{C} \mathcal{C}^T A^T \mathcal{B}^T \mathcal{B}\right] \\ &- \operatorname{trace}\left[A \left(\mathcal{B}^T \mathcal{D} \mathcal{C}^T\right)^T\right] \\ &+ \frac{1}{2} \operatorname{trace}\left[\mathcal{D} \mathcal{D}^T\right] \end{aligned} \tag{23}$$

The objective function is quadratic in A except for the following special cases: When $\mathcal{C}\mathcal{C}^T = I_{3\times3}$,

$$\begin{aligned} J(A) &= \frac{1}{2} \|\mathcal{B} A - \mathcal{D}\|_F^2 \\ &= - \operatorname{trace}\left[A \left(\mathcal{B}^T \mathcal{D}\right)^T\right] \\ &+ \frac{1}{2} \left(\operatorname{trace}\left[\mathcal{B}^T \mathcal{B}\right] + \operatorname{trace}\left[\mathcal{D} \mathcal{D}^T\right]\right) \end{aligned} \tag{24}$$

When $\mathcal{D}\mathcal{D}^T = I_{3\times3}$,

$$\begin{aligned} J(A) &= \frac{1}{2} \|A \mathcal{C} - \mathcal{D}\|_F^2 \\ &= - \operatorname{trace}\left[A \left(\mathcal{D} \mathcal{C}^T\right)^T\right] \\ &+ \frac{1}{2} \left(\operatorname{trace}\left[\mathcal{C} \mathcal{C}^T\right] + \operatorname{trace}\left[\mathcal{D} \mathcal{D}^T\right]\right) \end{aligned} \tag{25}$$

When $\mathcal{B}\mathcal{B}^T = \mathcal{C}\mathcal{C}^T = I_{3\times3}$,

$$\begin{aligned} J(A) &= \frac{1}{2} \|A - \mathcal{D}\|_F^2 \\ &= - \operatorname{trace}\left[A \mathcal{D}^T\right] + \frac{1}{2} \left(3 + \operatorname{trace}\left[\mathcal{D} \mathcal{D}^T\right]\right) \end{aligned} \tag{26}$$

Comparing these results with Eq. (10), we can see that they are equivalent to the Wahba problem and thus have closed-form solution as well with $B = \mathcal{B}^T \mathcal{D}$, $B = \mathcal{D} \mathcal{C}^T$, and $B = \mathcal{D}$, respectively. In general, however, the weighted orthogonal Procrustes problem does not have closed-form solution.

E. General Minimization Problem in Attitude Estimation

The minimization problems reviewed in the previous subsections can be written in a standard form, given by

$$J(A) = \frac{1}{2} \sum_{k=1}^{N} \|\mathcal{B}_k A \mathcal{C}_k - \mathcal{D}_k\|_F^2 \tag{27}$$

where N is the number of observations.

For the Wahba problem, even if W is not the identity matrix up to a scalar, the objective function can be rewritten as

$$J(A) = \frac{1}{2}\sum_{i=1}^{n_v} (\mathbf{b}_i - A\mathbf{r}_i)^T W_i (\mathbf{b}_i - A\mathbf{r}_i) = \frac{1}{2}\sum_{i=1}^{n_v} \left\| \sqrt{W_i}^T A\mathbf{r}_i - \sqrt{W_i}^T \mathbf{b}_i \right\|_F^2 \tag{28}$$

where $\sqrt{W_i}$ is the square root of W_i, satisfying

$$\sqrt{W_i}\sqrt{W_i}^T = W_i \tag{29}$$

The new form of the objective function is consistent with Eq. (27) with

$$\mathcal{B}_k = \sqrt{W_i}^T \tag{30a}$$
$$\mathcal{C}_k = \mathbf{r}_i \tag{30b}$$
$$\mathcal{D}_k = \sqrt{W_i}^T \mathbf{b}_i \tag{30c}$$
$$N = n_v \tag{30d}$$

If $W_i = W$, the objective function can be written in a more compact form, given by

$$J(A) = \frac{1}{2}\|\mathcal{B}A\mathcal{C} - \mathcal{D}\|_F^2$$

with

$$\mathcal{B} = \sqrt{W}^T \tag{31a}$$
$$\mathcal{C} = [\mathbf{r}_1, \cdots, \mathbf{r}_{n_v}] \tag{31b}$$
$$\mathcal{D} = \sqrt{W}^T [\mathbf{b}_1, \cdots, \mathbf{b}_{n_v}] \tag{31c}$$

For the GPS attitude determination problem, the objective function can be rewritten as

$$\frac{1}{2}\sum_{i=1}^{m_d}\sum_{j=1}^{n_d} w_{ij}\left[d_{ij} - (\mathbf{b}_i^d)^T A\mathbf{r}_j^d\right]^2 = \frac{1}{2}\sum_{i=1}^{m_d}\sum_{j=1}^{n_d} \left\|\sqrt{w_{ij}}(\mathbf{b}_i^d)^T A\mathbf{r}_j^d - \sqrt{w_{ij}}d_{ij}\right\|_F^2 \tag{32}$$

It is consistent with Eq. (27) with

$$\mathcal{B}_k = \sqrt{w_{ij}}(\mathbf{b}_i^d)^T \tag{33a}$$
$$\mathcal{C}_k = \mathbf{r}_i^d \tag{33b}$$
$$\mathcal{D}_k = \sqrt{w_{ij}}d_{ij} \tag{33c}$$
$$N = m_d\, n_d \tag{33d}$$

If the weights w_{ij} can be written as $w_{ij} = w_i w_j$, that is, there exists two vectors $\mathbf{w}_b$ and $\mathbf{w}_r$ that satisfy[7]

$$W^d = \begin{bmatrix} w_{11} & \cdots & w_{1n_d} \\ \vdots & \ddots & \vdots \\ w_{m_d1} & \cdots & w_{m_dn_d} \end{bmatrix} = \mathbf{w}_b \mathbf{w}_r^T \tag{34}$$

the objective function can be written in a more compact form, given by[7]

$$J(A) = \frac{1}{2}\|\mathcal{B}A\mathcal{C} - \mathcal{D}\|_F^2$$

with

$$\mathcal{B} = \sqrt{W_b}(B^d)^T \tag{35a}$$
$$\mathcal{C} = R^d\sqrt{W_r} \tag{35b}$$
$$\mathcal{D} = \sqrt{W_b}D\sqrt{W_r} \tag{35c}$$

where

$$W_b = \operatorname{diag}(\mathbf{w}_b) \tag{36a}$$
$$W_r = \operatorname{diag}(\mathbf{w}_r) \tag{36b}$$
$$B^d = \begin{bmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_{m_d}\end{bmatrix} \tag{36c}$$
$$R^d = \begin{bmatrix}\mathbf{r}_1 & \cdots & \mathbf{r}_{n_d}\end{bmatrix} \tag{36d}$$
$$D = \begin{bmatrix} d_{11} & \cdots & d_{1n_d} \\ \vdots & \ddots & \vdots \\ d_{m_d1} & \cdots & d_{m_dn_d}\end{bmatrix} \tag{36e}$$

III. Solution to Minimization Problems in Attitude Estimation

Expanding the general objective function in Eq. (27) gives

$$\begin{aligned} J(A) = &\frac{1}{2}\sum_{k=1}^{N} \operatorname{trace}\left[A\mathcal{C}_k\mathcal{C}_k^T A^T \mathcal{B}_k^T\mathcal{B}_k\right] \\ &-\sum_{k=1}^{N} \operatorname{trace}\left[A\left(\mathcal{B}_k^T\mathcal{D}_k\mathcal{C}_k^T\right)^T\right] \\ &+\frac{1}{2}\sum_{k=1}^{N} \operatorname{trace}\left[\mathcal{D}_k\mathcal{D}_k^T\right] \end{aligned} \tag{37}$$

Because the last term does not depend on A, the objective function is equivalent to

$$J(A) = \frac{1}{2}\sum_{k=1}^{N} \operatorname{trace}\left[A\mathcal{C}_k\mathcal{C}_k^T A^T \mathcal{B}_k^T\mathcal{B}_k\right] - \sum_{k=1}^{N} \operatorname{trace}\left[A\left(\mathcal{B}_k^T\mathcal{D}_k\mathcal{C}_k^T\right)^T\right] \tag{38}$$

which will be used in the development of the solution.

A. General Solution

The objective of the solution is to find the global minimum of the minimization problem. One issue needs to be handled in attitude estimation is attitude parameterization. Many attitude representations exist.[3] Three-parameter attitude representations are not used because the singularity issues cannot be avoided in global attitude estimation and must be handled with care. The general solution in this subsection is valid for both attitude quaternion and the attitude matrix.

The attitude is assumed to be parameterized by an $n \times 1$ vector $\mathbf{x}$. That is, $A = A(\mathbf{x})$. The goal is to find the optimal solution that minimizes the objective function, given by

$$J(\mathbf{x}) \triangleq J(A(\mathbf{x})) \tag{39}$$

subject to $m = n - 3$ equality constraints, given by

$$\mathbf{c}(\mathbf{x}) = \begin{bmatrix} c_1(\mathbf{x}) \\ \vdots \\ c_m(\mathbf{x}) \end{bmatrix} = \mathbf{0}_{m\times 1} \tag{40}$$

The method of Lagrange multiplier is used to solve the minimization problem. The $m \times 1$ vector of Lagrange multipliers is denoted by

$$\boldsymbol{\lambda} = \begin{bmatrix}\lambda_1 & \cdots & \lambda_m\end{bmatrix}^T \tag{41}$$

The $(2n-3)\times 1$ augmented state and the augmented objective function are

$$\mathbf{x}^a = \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} \tag{42}$$

and

$$J^a(\mathbf{x}^a) = J^a(\mathbf{x}, \boldsymbol{\lambda}) = J(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{c}(\mathbf{x}) = J(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j c_j(\mathbf{x}) \tag{43}$$

respectively.

The gradient of J^a with respect to $\mathbf{x}$ is

$$\mathbf{g}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\partial J^a(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}} = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} + \boldsymbol{\lambda}^T G_c(\mathbf{x}) = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} + \sum_{j=1}^{m} \lambda_j \frac{\partial c_j(\mathbf{x})}{\partial \mathbf{x}} \tag{44}$$

where

$$G_c(\mathbf{x}) \triangleq \frac{\partial \mathbf{c}^T(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial c_1(\mathbf{x})}{\partial \mathbf{x}} & \cdots & \frac{\partial c_m(\mathbf{x})}{\partial \mathbf{x}} \end{bmatrix} \tag{45}$$

For $\mathbf{x}^*$ to be a local minimizer, the necessary conditions are[10]

$$\mathbf{g}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}_{n\times 1} \tag{46a}$$

$$\mathbf{c}(\mathbf{x}^*) = \mathbf{0}_{m\times 1} \tag{46b}$$

Note that for the attitude quaternion and the attitude matrix, the necessary condition is a system of polynomial equations. The solutions $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ to the $m+n$ equations are stationary points. They are local minimizers when the 3×3 Hessian matrix evaluated at them are positive definite,[10] denoted by

$$H(\mathbf{x}^*, \boldsymbol{\lambda}^*) > 0 \tag{47}$$

The Hessian matrix is three-dimensional because the attitude only has three degrees of freedom. The 3×3 Hessian matrix is obtained by projection of the $n\times n$ Hessian matrix defined below onto the three-dimensional constraint subspace determined by the m equality constraints.

$$\bar{H}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \left.\frac{\partial \mathbf{g}(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^T}\right|_{\mathbf{x}^*, \boldsymbol{\lambda}^*} = \left.\frac{\partial^2 J^a(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x} \partial \mathbf{x}^T}\right|_{\mathbf{x}^*, \boldsymbol{\lambda}^*} = \left.\frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T}\right|_{\mathbf{x}^*} + \sum_{j=1}^{m} \lambda_j^* \left.\frac{\partial^2 c_j(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T}\right|_{\mathbf{x}^*} \tag{48}$$

The $n\times 3$ matrix $Q(\mathbf{x}^*)$ used to perform the projection satisfies

$$Q^T(\mathbf{x}^*) G_c(\mathbf{x}^*) = 0_{3\times m} \tag{49a}$$

$$Q^T(\mathbf{x}^*) Q(\mathbf{x}^*) = I_{3\times 3} \tag{49b}$$

After this matrix is obtained, the 3×3 Hessian matrix is given by

$$H(\mathbf{x}^*, \boldsymbol{\lambda}^*) = Q^T(\mathbf{x}^*) \cdot \bar{H}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \cdot Q(\mathbf{x}^*) \tag{50}$$

The projection matrix is not unique, however. As a result, the 3×3 Hessian matrix is not unique, either. Any two of these 3×3 Hessian matrices are related by a coordinate transformation.

The outline of the solution is as follows:

1. Find all the stationary points, that is, all the solutions to Eq. (46)
2. Find the global minimizer from the stationary points that minimizes $J(\mathbf{x})$
3. Find the local minimizers whose Hessian matrices (computed using Eq. (50)) are positive definite

The solution differs from gradient-based solutions in that it is guaranteed to find the global minimum. The first step is the most important and only nontrivial step. The homotopy continuation method is used to find all the solutions to Eq. (46), which are $2n-3$ equations of $2n-3$ unknowns. In particular, a numerical tool called Hom4PS 2.0 is chosen for its high speed.[13] Note that the homotopy continuation based tools can only find all the complex solutions of Eq. (46). After the complex solutions are obtained, post-processing is used to retain only real solutions.

B. Quaternion Based Solution

In the quaternion based solution, the attitude matrix is parameterized by the attitude quaternion. The state is chosen as

$$\mathbf{x} = \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \tag{51}$$

The quaternion is subject to one equality constraint, given by

$$c(\mathbf{q}) = \frac{1}{2}(\mathbf{q}^T\mathbf{q} - 1) = 0 \tag{52}$$

and the gradient of the constraint is

$$\frac{\partial c(\mathbf{q})}{\partial \mathbf{q}} = \mathbf{q} \tag{53}$$

Only one Lagrange multiplier is needed and the 5×1 augmented state is given by

$$\mathbf{x}^a = \begin{bmatrix} \mathbf{q} \\ \lambda \end{bmatrix} \tag{54}$$

The main tasks of the derivation is to find the partials used to compute the gradient and Hessian matrix.

To compute the gradient, the following relations are used:[16]

$$\frac{\partial \,\mathrm{trace}\left[\frac{1}{2}A\mathcal{C}_k\mathcal{C}_k^T A^T \mathcal{B}_k\mathcal{B}_k^T\right]}{\partial A} = \mathcal{B}_k^T\mathcal{B}_k A\mathcal{C}_k\mathcal{C}_k^T \tag{55a}$$

$$\frac{\partial \,\mathrm{trace}\left[-A\left(\mathcal{B}_k^T\mathcal{D}_k\mathcal{C}_k^T\right)^T\right]}{\partial A} = -\mathcal{B}_k^T\mathcal{D}_k\mathcal{C}_k^T \tag{55b}$$

$$\frac{\partial J(A)}{\partial A} = \sum_{k=1}^{N}\left(\mathcal{B}_k^T\mathcal{B}_k A\mathcal{C}_k\mathcal{C}_k^T - \mathcal{B}_k^T\mathcal{D}_k\mathcal{C}_k^T\right) \tag{56}$$

$$\frac{\partial A}{\partial q_1} = \begin{bmatrix} 2q_1 & 2q_2 & 2q_3 \\ 2q_2 & -2q_1 & 2q_4 \\ 2q_3 & -2q_4 & -2q_1 \end{bmatrix} \tag{57a}$$

$$\frac{\partial A}{\partial q_2} = \begin{bmatrix} -2q_2 & 2q_1 & -2q_4 \\ 2q_1 & 2q_2 & 2q_3 \\ 2q_4 & 2q_3 & -2q_2 \end{bmatrix} \tag{57b}$$

$$\frac{\partial A}{\partial q_3} = \begin{bmatrix} -2q_3 & 2q_4 & 2q_1 \\ -2q_4 & -2q_3 & 2q_2 \\ 2q_1 & 2q_2 & 2q_3 \end{bmatrix} \tag{57c}$$

$$\frac{\partial A}{\partial q_4} = \begin{bmatrix} 2q_4 & 2q_3 & -2q_2 \\ -2q_3 & 2q_4 & 2q_1 \\ 2q_2 & -2q_1 & 2q_4 \end{bmatrix} \tag{57d}$$

Note that $\partial A/\partial q_i$, $i = 1, 2, 3, 4$, are based on Eq. (16).

The gradient, which is cubic in $\mathbf{q}$, is then given by

$$\mathbf{g}(\mathbf{q},\lambda) = \begin{bmatrix} \mathrm{trace}\left[\left(\frac{\partial J(A)}{\partial A}\right)^T\left(\frac{\partial A}{\partial q_1}\right)\right] \\ \mathrm{trace}\left[\left(\frac{\partial J(A)}{\partial A}\right)^T\left(\frac{\partial A}{\partial q_2}\right)\right] \\ \mathrm{trace}\left[\left(\frac{\partial J(A)}{\partial A}\right)^T\left(\frac{\partial A}{\partial q_3}\right)\right] \\ \mathrm{trace}\left[\left(\frac{\partial J(A)}{\partial A}\right)^T\left(\frac{\partial A}{\partial q_4}\right)\right] \end{bmatrix} + \lambda\mathbf{q} \tag{58}$$

So, the necessary condition for $(\mathbf{q}^*, \boldsymbol{\lambda}^*)$ to be a local minimizer is

$$\mathbf{g}(\mathbf{q}^*, \lambda^*) = \mathbf{0}_{4\times1} \tag{59a}$$
$$(\mathbf{q}^*)^T\mathbf{q}^* = 1 \tag{59b}$$

The Q matrix corresponding to the quaternion constraint is

$$Q(\mathbf{q}) = \begin{bmatrix} q_4 & -q_3 & q_2 \\ q_3 & q_4 & -q_1 \\ -q_2 & q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix} \tag{60}$$

It can be verified that

$$Q^T(\mathbf{q}^*)\mathbf{q} = \mathbf{0}_{3\times1} \tag{61a}$$
$$Q^T(\mathbf{q}^*)\, Q(\mathbf{q}^*) = I_{3\times3} \tag{61b}$$

To compute the 3×3 Hessian matrix, we still need $\frac{\partial \mathbf{g}(\mathbf{q},\lambda)}{\partial \mathbf{q}^T}$ and $\frac{\partial^2 c(\mathbf{q})}{\partial \mathbf{q}\partial \mathbf{q}^T}$. From Eq. (53),

$$\frac{\partial^2 c(\mathbf{q})}{\partial \mathbf{q}\partial \mathbf{q}^T} = I_{4\times4} \tag{62}$$

No simple expression for $\frac{\partial \mathbf{g}(\mathbf{q},\lambda)}{\partial \mathbf{q}^T}$ exits. However, it is straightforward to derive it from Eq. (58) using a numerical or symbolic method. Finally, the 3×3 Hessian matrix, which is quadratic in $\mathbf{q}$, is given by

$$H(\mathbf{q}^*, \boldsymbol{\lambda}^*) = Q^T(\mathbf{q}^*) \left(\left. \frac{\partial \mathbf{g}(\mathbf{q},\lambda)}{\partial \mathbf{q}^T} \right|_{\mathbf{q}^*,\lambda^*} \right) Q(\mathbf{q}^*) + \lambda^* I_{3\times3} \tag{63}$$

C. Attitude Matrix Based Solution

This solution solves for the elements of the attitude matrix directly. First, the following quantities are defined:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{64}$$

$$\mathbf{x} = \mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} \tag{65}$$

The attitude matrix has six constraints, given by

$$\mathbf{c}(\mathbf{a}) = \begin{bmatrix} \frac{1}{2}(\mathbf{a}_1^T\mathbf{a}_1 - 1) \\ \frac{1}{2}(\mathbf{a}_2^T\mathbf{a}_2 - 1) \\ \frac{1}{2}(\mathbf{a}_3^T\mathbf{a}_3 - 1) \\ \mathbf{a}_1^T\mathbf{a}_2 \\ \mathbf{a}_1^T\mathbf{a}_3 \\ \mathbf{a}_2^T\mathbf{a}_3 \end{bmatrix} = \mathbf{0}_{6\times1} \tag{66}$$

which is equivalent to $A^TA - I_{3\times3}$. So we need six Lagrange multipliers, given by

$$\boldsymbol{\lambda}^T = \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_{12} & \lambda_{13} & \lambda_{23} \end{bmatrix} \tag{67}$$

The 15×1 augmented state and the augmented objective function are

$$\mathbf{x}^a = \begin{bmatrix} \mathbf{a} \\ \boldsymbol{\lambda} \end{bmatrix} \tag{68}$$

and

$$J^a(\mathbf{a}, \boldsymbol{\lambda}) = J(A(\mathbf{a})) + \boldsymbol{\lambda}^T \mathbf{c}(\mathbf{a}) \tag{69}$$

respectively.

To find the gradient, the objective function is first rewritten as

$$J^a(A, \Lambda) = J(A) + \frac{1}{2}\,\text{trace}\left[\Lambda(A^T A - 1)\right] \tag{70}$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & \lambda_{12} & \lambda_{13} \\ \lambda_{12} & \lambda_2 & \lambda_{23} \\ \lambda_{13} & \lambda_{23} & \lambda_3 \end{bmatrix} \tag{71}$$

From Eq. (55) and

$$\frac{\partial \frac{1}{2}\,\text{trace}\left[\Lambda(A^T A - 1)\right]}{\partial A} = A\Lambda \tag{72}$$

$$\frac{\partial J^a(A, \Lambda)}{\partial A} = \sum_{k=1}^{N} \left(\mathcal{B}_k^T \mathcal{B}_k A \mathcal{C}_k \mathcal{C}_k^T - \mathcal{B}_k^T \mathcal{D}_k \mathcal{C}_k^T\right) + A\Lambda \tag{73}$$

So, the necessary condition in matrix form is

$$\left.\frac{\partial J^a(A, \Lambda)}{\partial A}\right|_{A^*} = \sum_{k=1}^{N} \left(\mathcal{B}_k^T \mathcal{B}_k A^* \mathcal{C}_k \mathcal{C}_k^T - \mathcal{B}_k^T \mathcal{D}_k \mathcal{C}_k^T\right) + A^*\Lambda^* = 0_{3\times 3} \tag{74a}$$

$$(A^*)^T A^* = I_{3\times 3} \tag{74b}$$

The equivalent vector form of the necessary condition is

$$K(\boldsymbol{\lambda}^*)\mathbf{a}^* - \mathbf{b} = \mathbf{0}_{9\times 1} \tag{75a}$$

$$\mathbf{c}(\mathbf{a}^*) = \mathbf{0}_{6\times 1} \tag{75b}$$

where K can be easily derived by comparing Eqs. (74a) and (75a). Note that K does not depend on the elements of A. The necessary condition means

$$\frac{\partial J^a(\mathbf{a}, \boldsymbol{\lambda})}{\partial \mathbf{a}} = K(\boldsymbol{\lambda})\mathbf{a} \tag{76}$$

and therefore

$$\bar{H}(\mathbf{a}, \boldsymbol{\lambda}) = \frac{\partial^2 J^a(\mathbf{a}, \boldsymbol{\lambda})}{\partial \mathbf{a} \partial \mathbf{a}^T} = K(\boldsymbol{\lambda}) \tag{77}$$

It can be shown that the gradients of the six equality constraints are

$$G_c(\mathbf{a}) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} & \mathbf{a}_2 & \mathbf{0}_{3\times 1} & \mathbf{a}_1 & \mathbf{0}_{3\times 1} & \mathbf{a}_3 \\ \mathbf{0}_{3\times 1} & \mathbf{0}_{3\times 1} & \mathbf{a}_3 & \mathbf{0}_{3\times 1} & \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} \tag{78}$$

A projection matrix corresponding to $G_c(\mathbf{a})$ is

$$Q(\mathbf{a}^*) = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{0}_{3\times 1} & -\mathbf{a}_3^* & \mathbf{a}_2^* \\ \mathbf{a}_3^* & \mathbf{0}_{3\times 1} & -\mathbf{a}_1^* \\ -\mathbf{a}_2^* & \mathbf{a}_1^* & \mathbf{0}_{3\times 1} \end{bmatrix} \tag{79}$$

Finally, the 3×3 Hessian matrix is

$$H(\mathbf{a}^*, \boldsymbol{\lambda}^*) = Q^T(\mathbf{a}^*) K(\boldsymbol{\lambda}^*) Q(\mathbf{a}^*) \tag{80}$$

IV. Simulation Results

A Wahba problem and a GPS attitude determination problem are solved using HOM4PS 2.0 and the results are shown in this section. HOM4PS 2.0 is released as an executable, which loads the polynomial system by reading an input data file and saves the solutions to an output data files. Two MATLAB programs are written to prepare the input data file and process the output data file. All the programs are executed on a Macintosh computer with a 3 GHz Intel Core 2 Duo processor. The execution time of HOM4PS 2.0 for the Wahba problem is less than 0.5 seconds and that for the GPS problem is less than 1 second.

A. Wahba Problem

Since closed-form solutions to the Wahba problem exist, the purpose of this example is mainly to verify the proposed method. The true attitude quaternion is randomly chosen as

$$\mathbf{q} = \begin{bmatrix} 0.508975066874903 \\ 0.562504911614453 \\ 0.378089905856714 \\ 0.530641714152371 \end{bmatrix} \tag{81}$$

Two pairs of unit vectors observations are used, with $\mathbf{b}_1$ and $\mathbf{b}_2$ corrupted by zero-mean Gaussian noise with covariance $\sigma^2 I_{3\times 3}$. The vector representations and the standard deviation σ are:

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 0.081851273681315 & 0.746662261605013 \\ 0.171345905876038 & 0.259502682042531 \\ 0.981804944750365 & 0.612498020492817 \end{bmatrix} \tag{82a}$$

$$\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0.707106781186547 \\ 0 & 0.707106781186547 \\ 0 & 0 \end{bmatrix} \tag{82b}$$

$$\sigma_i = \sigma = 0.001 \tag{82c}$$

The attitude profile matrix and the K matrix that are used in the closed-form solution are

$$B\sigma^2 = \begin{bmatrix} 0.609821222118304 & 0.527969948436989 & 0 \\ 0.354842012084408 & 0.183496106208370 & 0 \\ 1.414906448504173 & 0.433101503753808 & 0 \end{bmatrix} \tag{83}$$

and

$$K\sigma^2 = \begin{bmatrix} 0.426325115909934 & 0.882811960521397 & 1.414906448504173 & -0.433101503753808 \\ 0.882811960521397 & -0.426325115909934 & 0.433101503753808 & 1.414906448504173 \\ 1.414906448504173 & 0.433101503753808 & -0.793317328326674 & 0.173127936352581 \\ -0.433101503753808 & 1.414906448504173 & 0.173127936352581 & 0.793317328326674 \end{bmatrix} \tag{84}$$

Table 1. Result of Wahba Problem

$(\mathbf{q}^*)^T$				$J(\mathbf{q}^*)\sigma^2$	$\Theta(\mathbf{q}^*)$ (°)
0.509216656365254	0.562250854442233	0.378006225214606	0.530738793813090	1.4998E-08	0.043
0.634995428434138	−0.146127271331926	0.324583278680832	−0.685618933257850	0.58596	179.989
0.146127271331925	0.634995428434138	−0.685618933257850	−0.324583278680832	3.4140	179.989
0.562250854442233	−0.509216656365255	−0.530738793813090	0.378006225214606	4.0000	179.960

The numerical result shows that Eq. (58) has eight solutions but only four distinct quaternion solutions. Two solutions are considered identical if one is $\mathbf{q}$ and the other is $-\mathbf{q}$. The solutions are listed in the first column of Table 1. The second column gives the values of the objective function, and the last column gives the attitude errors in degrees. The attitude error is defined as the minimum angle of rotation from the estimated attitude to the true attitude, which is the angle of rotation about the principal axis of rotation. The maximum possible attitude error is 180 degrees. The four quaternion solutions are identical to the four eigenvectors of the K matrix up to the sign. From the objective function, the first solution is the global minimizer.

The Hessian matrix for the first solution is

$$H\sigma^2 = \begin{bmatrix} 5.743374894802064 & -0.831144538481740 & -2.151043744762597 \\ -0.831144538481741 & 7.613181451309464 & -1.308781496466830 \\ -2.151043744762597 & -1.308781496466831 & 2.643443533901671 \end{bmatrix} \quad (85)$$

It can be verified that the matrix is positive definite. In fact, of the four Hessian matrices, this is the only one that is positive definite. The Hessian matrix of the fourth solution is negative definite, which means that that solution is the local maximizer (also the global maximizer). So, the Wahba problem has one local minimizer and one local maximizer.

Finally, it is worthy of mentioning that the above Hessian matrix is identical to a known result, given by

$$H = 2Q^T(\mathbf{q}_{max})(-K + \lambda_{max} I_{4\times 4})Q(\mathbf{q}_{max}) \quad (86)$$

where λ_{max} is the maximum eigenvalue of K and $\mathbf{q}_{max}$ is the eigenvector corresponding to λ_{max}, and the projection matrix is

$$Q(\mathbf{q}_{max}) = \begin{bmatrix} 0.530738793813090 & -0.378006225214606 & 0.562250854442233 \\ 0.378006225214606 & 0.530738793813090 & -0.509216656365254 \\ -0.562250854442233 & 0.509216656365254 & 0.530738793813090 \\ -0.509216656365254 & -0.562250854442233 & -0.378006225214606 \end{bmatrix} \quad (87)$$

B. GPS Attitude Determination

Two cases are tested. Case 1 uses three baselines and two GPS sightlines, with two baselines almost parallel to each other.[7] Case 2 uses two baselines and two GPS sightlines and its attitude solution is known to have ambiguity. The configurations are far from optimal, but the purpose is to test the algorithms rather than find the best configurations.

As the solutions of the Wahba problem, the quaternion solutions appear in pairs, each pair corresponding to one attitude. That is observed in all quaternion based solutions. The number of distinct solutions varies from run to run. The observed numbers of distinct solutions are 8, 10 and 12. With the baselines ($\mathbf{b}_i^d$) and the GPS sightlines ($\mathbf{r}_i^d$) fixed, the number heavily depends on the attitude, but no simple rule of the number of solutions is found. Regardless of the number of distinct quaternion solutions, there are always two solutions whose 3×3 Hessian matrices are positive definite and two solutions whose 3×3 Hessian matrices are negative definite. That is, there are always two local minimizers and two local maximizers. In the ensuing, typical individual results of the two cases are given.

1. Case 1

The randomly chosen quaternion is

$$\mathbf{q} = \begin{bmatrix} 0.494155327437997 \\ 0.577896222415969 \\ 0.583581779663523 \\ 0.285094326367176 \end{bmatrix} \quad (88)$$

The baselines, sightlines, direction cosine measurements (corrupted by zero-mean Gaussian noise with variance σ^2), and the standard deviation of the measurements are given by

$$\begin{bmatrix}\mathbf{b}_1^d & \mathbf{b}_2^d & \mathbf{b}_3^d\end{bmatrix} = \begin{bmatrix}0 & 0 & 0.099014754297667\\ 1 & 0 & 0.990147542976674\\ 0 & 1 & 0.099014754297667\end{bmatrix} \tag{89a}$$

$$\begin{bmatrix}\mathbf{r}_1^d & \mathbf{r}_2^d\end{bmatrix} = \begin{bmatrix}0.577350269189626 & 0\\ 0.577350269189626 & 0.707106781186547\\ 0.577350269189626 & 0.707106781186547\end{bmatrix} \tag{89b}$$

$$\begin{bmatrix}d_{11} & d_{12}\\ d_{21} & d_{22}\\ d_{31} & d_{32}\end{bmatrix} = \begin{bmatrix}0.591783588074708 & 0.555790541812144\\ 0.658944044778138 & 0.166861032642428\\ 0.696304319413962 & 0.648429425909754\end{bmatrix} \tag{89c}$$

$$\sigma_{ij} = \sigma = 0.001 \tag{89d}$$

The separation of the first and third baselines is approximately 8 degrees.

Table 2. Result of GPS Attitude Determination (Case 1)

$(\mathbf{q}^*)^T$				$J(\mathbf{q}^*)\sigma^2$	$\Theta(\mathbf{q}^*)$ (°)
0.494409741491392	0.577593314343100	0.583466310765854	0.285503125982629	6.9939E-07	0.067
0.023159988834038	0.545006738454282	−0.105182958951276	0.831485306606729	0.007896	119.726
−0.183182468102946	−0.245375932645723	−0.213643927625006	−0.927680498472893	0.13753	103.152
0.061334065674192	0.782910538873042	0.527985763278814	0.323295923705025	0.13987	55.977
−0.185595602113068	−0.607022044282350	−0.379301603783612	−0.673207845764400	0.15129	62.306
−0.275139858639560	−0.535339437079248	−0.260401946847975	0.754917592438580	0.28459	135.075
−0.170432201579888	0.349935356997199	0.699026105836855	−0.599883833709859	0.29167	138.423
0.097027598411826	0.431096166156030	0.543354233267374	−0.713798233300160	0.3574	131.506
0.409425112044050	−0.518623643018304	0.750344200570020	0.019600387778757	5.2727	139.504
0.880452553681176	−0.349323328370466	0.266764388871851	−0.177801219912304	5.2736	140.466
−0.697858607337799	0.464304407532680	−0.538523307847167	0.086066417467088	5.2796	137.029
−0.650162987939201	−0.235528668664488	0.669110759017677	0.272222569817222	5.5569	178.774

Twelve solutions (stationary points) are found in this run, summarized in Table 2. The first two solutions are the local minimizers. One differs from the true attitude by less than 0.1 degrees; the other differs from the true attitude by almost 120 degrees. Their Hessian matrices are

$$H\sigma^2 = \begin{bmatrix}5.947898283494723 & -2.926676996645922 & -3.048836914016578\\ -2.926676996645922 & 3.528598307001146 & -0.160789158652216\\ -3.048836914016579 & -0.160789158652215 & 6.410445709888654\end{bmatrix} \tag{90}$$

and

$$H\sigma^2 = \begin{bmatrix}6.571180128928024 & 2.584801773679908 & 2.997342203273436\\ 2.584801773679907 & 2.901044328250977 & -0.467657728186499\\ 2.997342203273436 & -0.467657728186499 & 6.211375418141446\end{bmatrix} \tag{91}$$

respectively. It can be verified that they are positive definite.

2. Case 2

In Case 2, the third baseline is removed and a different true attitude is used. All the other conditions remain unchanged. The data are shown below and the result of the 12 solutions (stationary points) in this sun is summarized in Table 3.

$$\mathbf{q} = \begin{bmatrix}0.867854445111753\\ 0.203020925783293\\ 0.452715585860023\\ 0.025685873808016\end{bmatrix} \tag{92}$$

$$\begin{bmatrix}\mathbf{b}_1^d & \mathbf{b}_2^d\end{bmatrix} = \begin{bmatrix}0 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} \tag{93a}$$

$$\begin{bmatrix}\mathbf{r}_1^d & \mathbf{r}_2^d\end{bmatrix} = \begin{bmatrix}0.577350269189626 & 0\\ 0.577350269189626 & 0.707106781186547\\ 0.577350269189626 & 0.707106781186547\end{bmatrix} \tag{93b}$$

$$\begin{bmatrix}d_{11} & d_{12}\\ d_{21} & d_{22}\end{bmatrix} = \begin{bmatrix}-0.208625565751983 & -0.483642936843990\\ 0.199040994199646 & -0.316026882228789\end{bmatrix} \tag{93c}$$

$$\sigma_{ij} = \sigma = 0.001 \tag{93d}$$

Table 3. Result of GPS Attitude Determination (Case 2)

$(\mathbf{q}^*)^T$				$J(\mathbf{q}^*)\sigma^2$	$\Theta(\mathbf{q}^*)$ (°)
0.867429762020386	0.202439020938956	0.453709587753502	0.027049228142041	5.909177458E-06	0.210
0.177675121707153	−0.632492159572748	0.594238774282549	−0.463967130672556	5.909177458E-06	147.134
−0.243945071722753	−0.495456850518871	0.478303456357552	0.682817043498858	0.34357	171.027
−0.688552516242612	z0.310329347315595	0.655319776219990	−0.012129281331835	0.34357	152.440
0.515487238280812	−0.405186760200873	0.754058746479743	−0.038626459648028	0.39062	90.262
0.819710358833898	−0.337191490253139	0.391817553353838	−0.246689747263808	0.39062	71.027
0.689822083186759	−0.377901820966324	0.597653924700860	−0.155388202730696	0.39221	75.905
0.661261214522559	0.290475532746040	−0.644689045103067	−0.250466776575558	0.66838	140.909
0.209680549253409	0.817837436785233	0.420055145633886	0.332760678091394	1.7395	113.715
0.281274555509038	0.383563870250663	0.087030793730653	0.875322239409086	1.7395	134.854
0.250466776575558	0.644689045103066	0.290475532746040	0.661261214522559	1.7485	120.431
0.155388202730698	−0.597653924700856	−0.377901820966323	0.689822083186762	2.0247	163.922

The first two solutions are the local minimizers. They correspond to two attitudes that are more than 100 degrees away from each other but yield the same value of the objective function, 0.000005909177458. This clearly shows the ambiguity of the two-baseline/two-sightline case.

V. Conclusions

A novel numerical method was presented for solving the minimization problem in attitude estimation. Based on the homotopy continuation method, which is used to find all the stationary points of the minimization problem, it is guaranteed to find the globally optimal solution to the minimization problem. The homotopy continuation method is suited to attitude estimation, which leads to a small-scale system of polynomial equations to solve. The method also helps to understand the properties of the GPS attitude determination problem. The numerical results show that the GPS attitude determination problem has two local minimizers and as many as 12 additional stationary points (including the two local minimizers).

References

[1] Wahba, G., "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, pp. 409.

[2] Shuster, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp. 79–88.

[3] Shuster, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, Oct.-Dec. 1993, pp. 439–517.

[4] Shuster, M. D. and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, 1981, pp. 70–77.

[5] Markley, F. L., "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *The Journal of the Astronautical Sciences*, Vol. 36, No. 3, July-September 1988, pp. 245–258.

[6] Markley, F. L. and Mortari, D., "Quaternion Attitude Estimation Using Vector Observations," *Journal of the Astronautical Sciences*, Vol. 48, No. 2/3, 2000, pp. 359–380.

[7]Crassidis, J. L. and Markley, F. L., "A New Algorithm for Attitude Determination Using Global Positioning System Signals," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, Sept.-Oct. 1997, pp. 891–896.

[8]Bar-Itzhack, I. Y., Garrick, J. C., and Montgomery, P. Y., "Algorithms for attitude determination using the Global Positioning System," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 6, November-December 1998, pp. 846–852.

[9]Nadler, A., Bar-Itzhack, I. Y., and Weiss, H., "Iterative Algorithms for Attitude Estimation Using Global Positioning System Phase Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 5, Sept.-Oct. 2001, pp. 983–990.

[10]Nocedal, J. and Wright, S. J., *Numerical Optimization*, Springer, New York, NY, 2nd ed., 2006.

[11]Morgan, A., *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*, Prentice Hall, Englewood Cliffs, NJ, 1987.

[12]Sommese, A. J. and Wampler, C. W., *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, River Edge, NJ, 2005.

[13]Lee, T. L., Li, T. Y., and Tsai, C. H., "Hom4PS-2.0, A Software Package for Solving Polynomial Systems by the Polyhedral Homotopy Continuation," *Computing*, Vol. 83, No. 109-133, 2008.

[14]Bates, D. J., Hauenstein, J. D., Sommese, A. J., and Wampler, C. W., "Bertini: Software for Numerical Algebraic Geometry," Available at http://www.nd.edu/∼sommese/bertini.

[15]Verschelde, J., "Algorithm 795: PHCpack: a General-Purpose Solver for Polynomial Systems," *ACM Transactions on Mathematical Software*, Vol. 25, No. 2, 1999, pp. 251–276.

[16]Crassidis, J. L. and Junkins, J. L., *Optimal Estimation of Dynamic Systems*, Boca Raton, FL: Chapman & Hall/CRC, 2004.

[17]Viklands, T., *Algorithms for the Weighted Orthogonal Procrustes Problem and Other Least Squares Problems*, Ph.D. thesis, Umeå University, Umeå, Sweden, 2006.